The influence of management style and quality on the operation of incentive pay: school teachers' performance pay in England and Wales 2014-2015

David Marsden, London School of Economics and Lisa Sezer, De Montfort University

Corresponding author David Marsden: d.marsden@lse.ac.uk

Draft: 7.Dec.2017

Paper submitted to the Labor and Employment Relations Conference, June 2018[1]

Contents

Acknowledgements: This paper has benefited from research assistance by Dr. Lisa Sezer, and advice and support from the stakeholder organisations in school teaching in England and Wales, and notably the teachers' unions which distributed the questionnaires and helped pilot its contents. It has benefited from a grant from LSE's STICERD foundation. The project extends LSE's Centre of Economic Performance research on performance pay. The views expressed in this paper are those of the author and do not necessarily reflect those of any of the above organisations. Acknowledgements are also due for use of the 2011 Workplace Employment Relations Survey results for benchmarking purposes.

Effects of management style and quality on the operation of incentive pay: school teachers' performance pay in England and Wales 2014-2015

1. The effect of management style and quality on performance related pay.

In recent years, there has been a growing research programme on performance related pay incentives for professional and highly qualified workers, and notably in the education sector. This is a group of workers for whom management has limited control of the work process because of the former's greater knowledge, and output is typically hard to measure in the short-term. They differ therefore from the windshield replacers of Lazear's (1996) famous study, and those of an earlier generation of studies of piecework (eg. Brown 1962, Schmiede and Schudlich, 1976). Typically, performance pay involves an element of agreement over objectives, and managerial judgement as to whether or not they have been achieved, and the cycle of objective setting and review often implies a longer-duration employment relationship. Given the central role of management in the setting and appraising of objectives, it is natural to ask how far the quality of management affects any potential motivational effects of performance pay. Often this is hard to investigate because of selection effects: that well-managed organisations may be more likely to adopt such practices. The new system for school teachers in England and Wales avoids this problem. Although schools have some flexibility over its implementation participation in the scheme and following its central provisions is compulsory for all state schools. Moreover, part of the brief of the national school inspectorate, Ofsted (England) and Estrin (Wales) is to monitor its operation in their school inspections.

Recent work by economists has advanced our knowledge considerably in recent years owing to improved data availability and use of panel and quasi-experimental methods. Existing research showing that the effectiveness of individual teachers and the quality of school leadership can improve pupil outcomes has been reviewed by Murphy (2011, 2012) and Husain (2013). There are considerable variations in teacher effectiveness in terms of pupil achievements, and especially for those from disadvantaged backgrounds (Murphy, 2013, Slater et al. 2012 for the UK, and for the US, Rivkin et al 2005, and Aaronson et al 2007). Moreover, differences in teachers' prior experience and educational qualifications appear to account for only a small part of differences in effectiveness. This strand of research emphasises two main policy instruments to improve teaching performance: incentives, and monitoring.

Incentives for teachers have been found to improve student outcomes in a variety of educational systems, although as Podgursky and Springer (2007) observe, the number of studies remains fairly small, and the design of incentive schemes is very diverse. Atkinson et al (2009) found that many teachers in England had responded to the financial incentives provided by the Threshold system introduced in 2000, and that this raised pupil performance for eligible teachers, although the School Teachers Pay Review Body questioned whether this effect was sustained in succeeding years (STRB 2012). In Israel, Lavy (2009) found that

tournament-type incentives for teachers in schools for disadvantaged pupils led to improved pupil performance. In India, Muralidharan and Sunsaraman (2011) and Duflo et al (2012) also using quasi-experimental methods found that financial incentives for teachers helped improve pupil performance by reducing absenteeism. In the US, Figlio and Kenny (2007) found that pupils achieved better test scores in schools with merit pay systems. However, these authors also commented that the available data sets lacked detailed information about school human resource practices.

Monitoring performance, how it is undertaken, on what it should focus, and its perceived fairness, is a key aspect of human resource policies. There has been concern that objective measures, such as value added in test scores, contain a lot of noise and fluctuate from year to year (Staiger and Rockoff, 2010), and they may bias performance in unintended ways, notably towards what can be measured (Murnane and Papay, 2010). This has led to increased research into the effects of subjective, or judgemental, measures. In the US, Rockoff and Speroni (2010) consider classroom observation by professional mentors and head teachers, and in England, Hussain (2013) considers judgements of school performance based on intensive observation of schools during Ofsted school inspections. Both studies find objective and judgemental measures are correlated but complementary as they convey different kinds of information about teaching and student progress. In particular, judgemental assessments by providing detailed feedback can improve performance by raising teaching skills (Taylor and Tyler, 2011). Reviewing the research from an economic perspective, Hanushek and Rivkin (2010) warn that despite the methodological quality of the individual research findings, concerns about accuracy, fairness, and potential adverse effects of incentives based on a limited set of outcomes raise worries about their use for education personnel and policy decisions.

Monitoring performance also raises the question of procedural justice in human resource policies. The process by which employees' objectives are set and their performance evaluated is critical, particularly where evaluation is dependent on line-manager judgements. It is widely argued that goal setting and appraisal systems, if they are to be effective, should be considered by employees to be fair and unbiased (Latham and Pinder, 2005), and to respect their own work values and desire for autonomy in their work (Bakker et al. 2007). This research also emphasises the importance of organisational support in the form of feedback, training and other resources to ensure employees have the means to achieve higher performance. Cropanzano et al (2007) argue that practices such as appraisal need to be governed by fair procedures (procedural justice), because if employees suspect management will conduct evaluations in bad faith, then they will adapt their behaviour accordingly, even when they have unequivocal evidence to the contrary. This will lead to greater resistance to management and a lesser willingness among employees to engage in 'organisational citizenship behaviour' (Shapiro and Kirkman, 2001). Erdogan et al (2001) find that fairness is needed both at the level of an appraisal system and its criteria, and at that of the behaviour of individual raters in such areas as feedback. Reviewing 35 years of empirical research on goal setting, Locke and Latham (2002) argue that employee involvement in goal setting, as opposed to 'top down' approaches, can benefit motivation not only by ensuring better

informed and more realistic goals, but also by encouraging employees to take greater ownership of their goals. Recent research has emphasised the role of appraisal in motivating better performance by fostering employee engagement (Gruman and Saks, 2011). This builds on a longer standing theme in the literature on employees' desires for autonomy and self-determination in their work, and 'voice' to influence their objectives, something likely to be important among teachers (Gagné and Forest, 2008, Cawley et al., 2001). The importance of the HR considerations is enhanced by the work of Kahneman and Tversky (2000) which shows that most people are twice as averse to potential losses as they are motivated by gains. The atmosphere in which performance management is operated can affect whether employees perceive it as taking something away from them, substituting conditional for automatic pay progression, or providing greater rewards and recognition for good performers.

Thus, the HR management research complements that of the economists, suggesting that if linking pay to performance is to motivate highly skilled professionals such as teachers, then one has to look closely at how well the establishment of objectives and evaluation procedures are operated and correspond to their aspirations. This paper focuses on employee motivation, that is, their willingness to perform. It explores how this is influenced by uneven access to performance pay increments, how well they believe managers run the appraisal and objective setting procedures in their schools, how far they feel management involves them in the establishment of the school-wide objectives into which their personal objectives are meant to be embedded, and at the general quality of management in their schools as evaluated by school inspectors.

This paper seeks to advance your understanding by drawing on both research streams considering both the incentive aspect of performance pay progression, and the quality of school leadership as seen in part by individual teachers, but also as assessed by school inspectors. It starts with a short presentation of key features of the performance pay scheme, followed by a brief description of the survey and methods. It then presents a descriptive account of teachers' perceptions of the scheme's motivation and divisiveness before turning to an analysis of the effects of some key factors on these perceptions. The first is the effect of teachers' position on their pay spine, as this determines their access to pay advancement. Comparing teachers who are eligible for performance pay progression with those who are not provides an indication of the salience of performance pay for teachers, and thus to its potential incentive effect, This serves also as a control variable for subsequent analysis of the link with appraisal quality, involvement, as seen by teachers themselves, and of leadership quality as seen by school inspectors. Finally, the paper examines whether teachers who said performance pay motivated them in 2014 were more likely to achieve their objectives in 2015.

2. The Performance pay system for teachers in England and Wales

The new system of performance pay for school teachers in England and Wales was introduced in state primary and secondary schools in 2013, and set out to revitalise and extend an earlier scheme introduced in 2000 (see DfE, 2014). First performance payments under the new scheme were made in the summer of 2014.

The scheme comprises three main elements: (i) to replace the former six annual pay increments for teachers on the 'Main Scale' by performance-based progression; (ii) to restore the performance element in progression at the 'Threshold' between the Main and the Upper pay scales, and (iii) to reinvigorate the performance element in progression on the three points on the Upper pay scale. The School Teachers Pay Review Body judged that the latter two had both lost their bite since their introduction in 2000 (STRB 2012).[2]

Performance evaluation begins by agreeing objectives at the start of the school year, mid-year feedback, and final assessment against the agreed objectives and national teacher standards at the end of the school year. This assessment should comprise a recommendation on pay which has to be reviewed by the school's senior management, and endorsed by the school's governing body. In their triennial school visitations, national school inspectors from the Office for Standards in Education, Ofsted, scrutinise whether performance awards are tracking overall pupil performance in the school.

3. The Survey

This study comprises a mixture of data from a survey of individual teachers which is linked to school-level data from the Education Department, and notably school inspection reports. For the survey, at the outset of the study, wide consultation was undertaken with all the main stakeholders, including the teachers' and head teachers' associations, the national representatives of school governors, of the local government employers in England and Wales, and with the Education Department. The stakeholders provided advice on the questionnaire design, and the unions also enabled it to be tested in focus groups of their members. The questionnaire was administered by the unions sending emails to a sample of their membership inviting them to reply via a link to an online confidential questionnaire held at a secure location. The survey has been funded by the STICERD research centre at the London School of Economics, and is part of the continuing independent research on performance pay by LSE's Centre for Economic Performance.[3]

So far, four surveys have been carried out, in the spring of 2014,, before the scheme came into effect and annually thereafter. This paper focuses on the results of the first two waves. In 2014, 4,372 usable replies were received, and in 2014, 1,306, and the panel element comprises about 750 individuals. The main focus in this paper is on the panel respondents because they enable one to compare replies from the *same individuals* in both years. In this paper, these results are linked to school level data from the Education Department and the school inspectorate's evaluation of school leadership..

As with all voluntary surveys, there is the possibility of bias arising from those who feel most strongly about the topic being the most willing to put aside their free time to respond. This is explored in more detail in Appendix 2 (available on request). Nevertheless, the sample response shows a good coverage of the general demographic characteristics of teachers, and by school type from the School Workforce Census. Some of the key attitudinal findings were also compared with those of the 2011 Workplace Employment Relations Survey on teachers' views about relationships in their schools shows that the current sample for 2014 is only

moderately more negative than in WERS, but this may also reflect a real change in attitudes owing to the uncertainties attached to the new pay system, and widespread concerns about budgets and teachers' workloads. There is some tendency for more senior teachers to respond, and as a result, most of the tables showing percentage distributions have been weighted to take account of this.

4. Teachers' assessment of performance pay and changes after one year of operation

In jobs that require employees to exercise a good deal of judgement and discretion in how they are carried out, management has to rely heavily upon their motivation. Although many teachers may feel that government targets and monitoring have increased greatly in recent decades, the nature of their tasks remains far more complex than administering bureaucratic routines or manual tasks that can be easily measured. Schools rely on teachers to work conscientiously with their students, constantly adapting to their educational and disciplinary needs. Hence, they rely heavily on teacher motivation. As motivation is a psychological state, the natural starting point when investigating the effects of performance pay schemes for this type of job is to enquire whether employees perceive them as motivating, relevant to their work, and fairly operated.

The motivational questions were divided roughly equally between those focusing on potential strengths and those focusing on potential weaknesses of performance pay for teachers. Thus among the first category were factors relating to its perceived motivation, including views about the principle of linking their pay to performance, whether doing so meant that good teaching is properly rewarded and is fairer, whether it helped them focus more on the objectives established in the performance review, whether it gave them an incentive to sustain or improve teaching quality, whether ssessment should, as it does, include elements of pupil progress, and whether it provides and incentive to remain a teacher. Among the second category, questions focused on factors that might contribute to perceived divisiveness, including potential favouritism, resentment between those who get awards and those who do not, whether it is seen primarily as a device to increase work intensity, whether it conflicts with team work in schools, whether the size of payments is too small to provide an incentive, whether teachers consider their work is already at an appropriate standard, and whether their school can afford to pay an award even if it is merited. Such factors could contribute to perceived divisiveness because they distort the distribution of awards, and prevent the scheme from meeting its stated aim of incentivising better teaching. The theoretical rationale behind these questions is provided by expectancy theory (Lawler 1971), which proposes that employees will respond to incentives if they value them (valence), if they have scope within their jobs to increase performance (instrumentality), and if they believe that management will measure and reward it competently and fairly (expectancy). This can be seen as a virtuous circle of positive motivation, but breaks in the linkages can also lead to a downward cycle leading to perceived divisiveness. This framework is congruent with the theory of incentives in mainstream personnel economics (eg. Lazear 2007), and so provides a convenient

approach to measuring employee perceptions of motivation and divisiveness. These questions were trialed on focus groups of teachers.

The responses to these two sets of questions are summarised in Tables 1 and 2 respectively. Both tables show how teachers' responses spread across the scale from strongly disagree to strongly agree with the statements in the questionnaire, the number of respondents in the panel in each year. As all sample surveys involve a margin of statistical error, the table also includes an indication of whether the difference shown between the two years might have come about by chance. Thus, for the first question in Table 1, about the principle of linking pay to performance, one can see that between 2014 and 2015 teachers have become more negative, and this change of this size is unlikely to have come about by chance (less than 1% probability (0.001).

Responses to all of the questions in Table 1 tend towards the negative end of the spectrum in both years, indicating that most teachers do not find PRP motivating: most remain opposed to the principle of linking pay progression to performance, they don't believe it means that good teaching is properly rewarded, or that it will result in a fairer allocation of pay, nor do most believe that it provides an incentive either to sustain or improve their quality teaching, to take performance review more seriously, or to remain in teaching. That said, on certain questions, around one fifth of teachers are more positive than their colleagues, notably on the principle, on reward for good teaching, on performance reviews, and on taking account of pupil performance in teachers' evaluation.

On the more negative aspects summarised in Table 2, teachers' responses also tend towards the negative side: most find it divisive. Consistently, between 70% and 80% of respondents suspect awards will be based on favouritism, that giving awards will create resentment among those who do not receive them, that the basic intention is to increase workloads, that the incentive is too small, that teachers already work at an appropriate standard, that its individual focus conflicts with team work in schools, and that their schools cannot afford to pay for performance even when it is merited. Although it may be suspected that discontented teachers responded disproportionately to the survey, these results are consistent with those of earlier surveys of classroom teachers and to a lesser degree of head teachers by the author in 1996/7 and in 2000/04 (Marsden and French, 1998, and Marsden and Belfield, 2000).

New pay systems introduce uncertainty into workplaces, and so one might expect initial reactions of suspicion and hostility before they bed down. Employees and managers who have become used to how the previous system worked, and who have developed a good working relationship under its auspices, may find these patterns of work challenged by major changes in the pay rules. Norms of fairness often also play an important part in pay, and these too can be disrupted by new pay systems: for example, people may accept seniority pay because they know their turn will come, but when the rules are changed they may feel cheated if their expected progression is jeopardised. Often too the weaker party in a bargaining relationship may fear that management will use the new rules to impose changes unilaterally, and use the increased uncertainty of the new system to its advantage, for example, by making pay progression depend upon employees accepting changes against their

will. Often too the uncertainty can generate a degree of suspicion about management's intentions which may subsequently prove to be unfounded. Hence, there are many reasons why one might expect initially negative employee views about a new pay system to moderate once they experience its operation in practice.

Table 1: Potentially motivating aspects of linking teachers' pay progression to performance: panel responses (weighted)

	Disagree strongly	Disagree	Hard to say	Agree	Agree strongly	Total %	N	Prob Chi2	
Is a good principle									
2014	32.0	24.4	15.6	25.6	2.5	100	784		
2015	38.6	23.7	18.5	16.9	2.3	100	745	*0.001*	****
It means that good teaching is properly rewarded									
2014	33.9	28.6	14.5	20.1	3.0	100	785		
2015	41.0	27.4	15.5	13.8	2.3	100	746	*0.005*	****
It will result in a fairer allocation of pay									
2014	41.6	39.7	11.8	6.2	0.7	100	782		
2015	47.6	32.5	15.2	3.8	0.9	100	745	*0.004*	****
It will make me take the objectives of my performance review more seriously									
2014	28.4	23.5	23.0	21.9	3.2	100	784		
2015	35.7	32.8	12.5	15.6	3.4	100	743	*0.000*	****
It will give me a real incentive to improve/sustain the quality of my teaching									
2014	42.2	32.1	19.4	5.5	0.9	100	784		
2015	46.5	35.7	9.9	7.0	0.9	100	745	*0.000*	****
It is good that individual teachers' pay should take some account of pupil performance									
2014	27.3	28.2	22.9	20.2	1.4	100	784		
2015	27.7	33.2	20.7	17.9	0.6	100	743	*0.129*	+
It makes it more attractive for me to remain a teacher									
2014	53.6	26.8	14.5	4.5	0.7	100.0	784		
2015	56.0	30.0	9.2	3.5	1.3	100.0	745	*0.015*	***

Note: statistical significance: significant at <1% ****; at <2% *** ; at <5%, ** ; <10%, * ; <20% + . A Chi 2 probability greater than 10% (0.1) signifies that there was probably no change in teachers' views between 2014 and 2015. Panel sample weighted to take account of higher response rate among more senior teachers using School Workforce Census 2013 distribution by spine point.

This benign expectation does not appear to be justified, at least for the sample of teachers in the panel survey. Their scepticism about the principle and the purported benefits of the new system appears if anything to have intensified with the experience of its first year of operation. On the principle, and on fairness, the share of negative and doubtful ('hard to say') views appears to have increased. On reinforcing the performance review and on incentives to sustain or improve teaching views have become more polarised as the 'hard to says' have

declined sharply, moving towards negative. In contrast, the question on whether it provides an incentive to remain a teacher, which was overwhelmingly negative before the new scheme came into operation, remains unchanged over the period.

Table 2: Potentially divisive aspects of linking teachers' pay progression to performance: panel responses

	Disagree strongly	Disagree	Hard to say	Agree	Agree strongly	Total %	N	Prob Chi2	
Leaders will use performance pay to reward their favourites									
2014	2.7	1.9	4.0	37.6	53.8	100	784		
2015	2.7	0.6	5.3	40.9	50.5	100	746	0.0634	*
It will cause resentment among teachers who feel they perform well but do not receive an award									
2014	2.4	6.4	20.6	32.5	38.1	100	785		
2015	2.8	7.1	24.7	31.8	33.6	100	745	0.2561	-
For all that is said about improving teaching quality, the new pay system is simply a device to get more work done									
2014	2.4	4.7	23.7	34.7	34.5	100	785		
2015	2.7	5.8	19.2	36.8	35.6	100	746	0.3179	-
The link is problematic because it is hard to relate the work done in schools to individual performance									
2014	0.9	2.2	7.1	31.9	58.0	100	784		
2015	1.6	1.9	6.2	36.0	54.3	100	745	0.272	-
It will have no effect on the quality of my work because it is already at the appropriate standard									
2014	2.1	4.7	18.7	35.4	39.2	100	784		
2015	2.4	4.4	17.9	36.9	38.4	100	743	0.953	-
The size of payments is too small to make me want to work harder to get them									
2014	4.0	10.9	46.1	25.2	13.8	100	785		
2015	5.5	12.4	32.6	29.9	19.6	100	742	0.000	****
Even if my performance is good enough, I doubt if my school can afford to reward me with a pay rise									
2014	2.2	5.3	22.0	33.6	36.9	100	785		
2015	2.3	8.0	21.2	27.3	41.2	100	745	0.0384	**

Panel sample, weighted to take account of higher response rate among more senior teachers using School Workforce Census 2013 distribution by spine point.

Turning to the 'divisiveness' questions (Table 2), about perceived favouritism, resentments, and the perceived mismatch between the scheme and teachers' views about the nature of their work, the picture was negative in 2014, and experience of the first year of operation of the scheme does not appear to have brought about any significant change in teachers' views. The exceptions concern a slight decline in perceived favouritism, and increased scepticism about the amount of money available as an incentive, and whether the school can afford to award performance pay rises. The latter two may reflect the general conditions of financial austerity during the scheme's implementation.

These results are general, across all teachers in the panel sample, are also consistent with those of the full sample which was compared separately with the panel results. However, not all teachers have the same opportunity to gain performance increments, such as those at the top of their respective pay scales, and this could affect perceptions of motivation and fairness. Two questions emerge: performance related pay progression was introduced at a time of severe public sector pay restraint, so that the opinions expressed could reflect general discontent with pay; and how far dissatisfaction is about the spread of a performance culture in schools, of which performance management is a central part, rather than about pay itself. The next section explores this by examining how teachers' pay scale position affects their views on whether PRP motivates.

5. The influence of teachers' position on the pay-spine on views of performance pay

With complex reward systems, such as when pay progression is conditional on a process of objective setting and performance appraisal, employees' judgements of them may be holistic, bundling together views about pay and appraisal, and it is possible that views about appraisal may crowd out those of incentive pay. One way to gauge the salience of pay as opposed to appraisal is to examine the experience of teachers who are subject to appraisal, but whose pay will not be affected by its outcome. This can be done by comparing the experience of those who were at the top of their pay scale in 2014 when the new scheme came into effect with those on the rungs below.

According to the School Workforce Census, in 2012/13, 35% of teachers were on the six point Main Scale, below its maximum, M6, and under the old scheme, would have been eligible for automatic annual pay progression. The same source also shows that around 40% of all classroom teachers were at the top of their respective pay scales, 13% at the top of the Main Scale, and 26% at the top of the Upper Scale. [4] Using the survey returns, a pay progression transition matrix for 2014-2015 reflects these percentages on their respective maxima, but also shows that a proportion of teachers at other points who might formerly have progressed did not do so. A fifth of those on M5 did not progress to M6, and about half of those on M6, and eligible to progress through the Threshold to the Upper Pay Scale stayed put. On the two lower rungs of the Upper Scale, about a third stayed put, although it should be said that progression there was originally designed to be performance related, although the STRB was skeptical about this in its 2012 report. Thus pay could be salient at different points on the pay scales: for those on M5, replacement of seniority by performance could be experienced as the loss of an entitlement. For those on M6, and ready for Threshold assessment and progression, the tightening up on assessments could also be experienced as a loss. On the other hand, teachers who progressed would not have experienced loss of anticipated pay increments, at least for the time being, and may even appreciate the opportunity.

We now turn to the effects of scope for pay progression at different positions on the pay spine on perceived motivation and divisiveness. The survey results enable the linking of information of teachers' current position on their pay scales to perceptions of motivation and

divisiveness of performance pay (Table 3). To explore these, the questions in Tables 1 and 2 were combined into indexes of positive motivation and divisiveness using factor analysis. In effect, these indexes show where an individual teacher's responses stand in relation to the average for all teachers in the sample: that is whether their individual perceptions of motivation or divisiveness of PRP are above or below the average. [5] As can be seen from Tables 1 and 2, the 'average view' is that PRP is not very motivating, and it is somewhat divisive. Views on PRP were then regressed on whether teachers had progressed up a spine point in 2014-15, and whether they had been on the maximum for their pay spine in 2014. Because opportunities for progression may vary for part-time teachers, and because there may be gender differences in attitudes, and because relations may differ between primary and secondary schools, the former often being smaller and more informal, these variables were also included. The summary and more detailed results for the panel respondents are shown in the two panels in Table 3.

In the left-hand panel, the most striking point is that receiving a performance pay progression in is associated with a more positive than average view of performance pay as a motivator: receiving an award is associated with teachers being 26% (0.263) more likely to be more positive than average about performance pay. On the other hand, being at the top of their pay spine, and thus ineligible for an award, means that they are about 18% more negative than average about its motivational aspects (-0.175). The same regression done for divisiveness, not displayed here, shows that being on the spine max increases perceived divisiveness about 20%.

The survey data also enable a more detailed look at individual spine points, and thus make it possible to disentangle the effects of different aspects of performance pay in schools (see the right-hand panel). This panel further distinguishes those who received pay progression in 2015 from those who did not. Thus, teachers on M5 and M6 in 2015 who had not progressed find performance pay less motivating than average (M6: -0.350, and U3: -0.191). On the intervening spine positions, apart from M5, there was no particular influence on perceived motivation.[6] This suggests that being on one's spine maximum contributes to the view that PRP is not motivating. In contrast, those who had progressed to these positions in 2015 were much more positive about performance pay: those who just progressed to M6 were 50% (0.526) more positive than average, and those on U1, having just passed the Threshold, were 71% (0.709) more so.

To judge from the differential relationships between pay spine position and teachers' perceptions of whether the new system is motivating or not, pay is clearly a salient factor affecting both negative and positive views. It is also clear that these judgements are much more than general dissatisfaction with pay restraint. Such dissatisfaction is most intense at the points where the incentive elements of the scheme bite sharpest.

The next section explores two potential channels through which the nature of the objective setting and appraisal process may influence perceptions of motivation and divisiveness. They concern whether appraisals are seen as being conducted in a systematic fashion, and whether

teachers feel involved in determining the overall objectives set out in the School Development Plan which is designed to play a central role in the objective setting process.

Table 3: Teachers' views of performance pay according to movement on their pay spine: 2015[7]

PRP motivates	Coef.	Std. Err.	Sig	Coef.	Std. Err.	Sig
Movement up a spine point 2014-2015	0.263	0.081	****			
On spine max in 2015	-0.175	0.072	***			
Spine Point in 2015						
M3-M4				0.120	0.176	-
M5				-0.970	0.117	****
M6				-0.350	0.163	**
U1				-0.078	0.194	-
U2				0.110	0.192	-
U3				-0.191	0.115	*
Leadership scale				0.231	0.326	-
Progression 2014-2015 to:						
M3_4				0.271	0.259	-
M5				0.979	0.189	****
M6				0.526	0.220	***
U1 (passing Threshold)				0.709	0.283	***
U2				-0.073	0.197	-
U3				0.065	0.123	-
Leadership scale				0.435	0.405	-
Part-time	-0.199	0.085	***	-0.169	0.088	*
Female	0.036	0.080	-	0.048	0.082	-
Primary school	0.053	0.084	-	0.029	0.083	-
Constant	-0.072	0.085	-	-0.056	0.111	-
Number of obs	790			790		
Prob > F	0.000		****	0.000		****
R-squared	0.0378			0.0868		

Notes: Regression weighted according to 2013 SWFC distribution of teachers by salary point. Spine point effects take junior teachers as the benchmark (unqualified and on points M1 and M2).[8]

Notes: OLS regression. Panel sample. Weighted to adjust for over-response in higher spine points using the SWFC 2013 distribution. Statistical significance: significant at <1% ****; at <2% *** ; at <5%, ** ; <10%, * ; <20% + . Indicators of PRP motivates and divisiveness derived using factor analysis based on questions in Tables 1 and 2, and missing observations for 2015 were assigned the value of the previous year. Movement up a spine point was calculated by comparing teachers' positions on their pay spine in 2014 with an estimate of their position in 2015. The same regressions for perceived divisiveness show it increased and was strongly significant with respect to being on the maximum of the teacher's pay spine (coefficient: 0.180), other coefficients in the left-hand panel of the table were not statistically significant. On the more detailed breakdown in the right hand

panel, divisiveness was significantly lower among teachers on M3_4 (-0.33), and among those who had just passed the threshold onto U1 (-0.47). All other variables were non-significant.

6. The influence of appraisal procedures and participative leadership

Existing research findings on whether employees find PRP motivating or divisive suggests that the quality of appraisal procedures and how far managers involve them in shaping performance objectives play an important role. Employees are more likely to find performance pay motivating if they believe evaluations of their performance are accurate and fair (Lawler, 1971, Latham and Locke, 2002), and they are more likely to take ownership of their objectives if they are involved in their formulation (Cawley et al, 1998).

Table 4: Objectives and feedback in performance reviews[9]

Objective setting							**Feedback**				
Review established specific objectives?							Did review give feedback? (a)				
	No	Yes, to some extent	Yes, defin-itely	Total %	N		No	Yes, to some extent	Yes, defin-itely	Total %	N
2014	1.2	18.5	80.3	100	670		23.0	23.3	53.7	100	789
2015	1.2	20.8	78.1	100	656		16.0	19.7	64.3	100	776
Prob Ch2				0.653	-					0.000	****
Know how objectives will be monitored?							Feedback: give reasons?				
2014	8.5	40.2	51.3	100	669		13.7	52.2	34.1	100	484
2015	7.5	40.0	52.4	100	656		11.9	46.8	41.3	100	527
Prob Chi2				0.837	-					0.075	*
Objectives include indicators of pupil progress?							Feedback: linked to performance against previous objectives?				
2014	4.1	11.7	84.2	100	671		8.6	36.8	54.7	100	486
2015	5.2	11.7	83.1	100	655		9.0	38.7	52.3	100	527
Prob Chi2				0.616	-					0.775	-
Opportunity to discuss objectives in review?							Feedback referred to classroom observation?				
2014	13.0	30.5	56.6	100.0	670		17.9	30.3	51.8	100	485
2015	12.5	34.7	52.8	100.0	655		14.3	38.1	47.6	100	527
Prob Chi2				0.287	-					0.022	**
Know how your appraisal will be linked to pay?							Feedback: include recommendation on pay?				
2014	40.5	34.5	25.0	100	667		71.2	8.5	20.3	100	466
2015	36.8	27.9	35.3	100	655		48.6	12.1	39.4	100	527
Prob Chi2				0.000	****					0.000	****

Note (a) The response categories for the first question were: no, not applicable, yes. Panel sample, weighted.

Prob chi 2: a figure greater than 0.1 signifies a more than 10% chance that the changes shown between 2014 and 2015 are due to statistical error. In other words, we can infer that there was no actual change in teachers' views.

To explore the link between attitudes to PRP and appraisal quality and participatory objective setting the most factual of the survey's questions relating to the latter were chosen. This is to minimise the risk that attitudes to PRP will colour teachers' perceptions of appraisal and involvement: for example, hostility to PRP could lead teachers to question the fairness of appraisal. Thus appraisal quality here refers primarily to teachers' views as to how *systematically* performance reviews are operated. As a first step, Table 4 reports teachers' reports of their experience of objective setting and feedback in appraisals, focusing on whether aspects of their reviews which are specified in Education Department guidelines have been carried out in their case (see DfE 2012 and 2014). To facilitate comparisons before and after introduction of the new link with pay, Table 4 results are confined to the same teachers who responded in both 2014 and 2015: the panel sample.

Procedures for objective setting appear somewhat better embedded than those for feedback. Almost all reviews in 2014 established specific objectives for the coming year, leaving only modest scope for improvement in 2015. In contrast, feedback appears to have been less well embedded in 2014, and to have shown some improvement by 2015: the proportion in the panel reporting no feedback fell from 23% to 16%.[10] Several of the other questions revealed little change with the introduction of a more systematic link with pay. The big exception concerns an increase in the proportion of teachers who definitely know how their objectives will be linked with pay (up from 25% to 35%), and receiving a recommendation on pay in their performance review (up from 20% to 39%).

To explore the link between the quality of performance review procedures and experiences with performance pay, the questions in Table 4 were condensed into an index using factor analysis based on the full sample for both years, similar to that used for whether PRP was seen as motivating or divisive.[11] Teachers' replies were then divided according to whether their factor scores were above or below the mean. Thus, in Table 5, in 2015, better quality appraisal procedures were associated with a greater proportion of teachers reporting relatively positive attitudes to PRP: 48% compared with 42%. Likewise, in 2015, more systematic appraisal procedures were associated with less teachers finding PRP divisive (42% compared with 56%), more systematic appraisals being those that conformed more closely to official guidelines. It appears also that the relationship between use of systematic procedures for appraisal and positive attitudes to PRP has strengthened between 2014 and 2014, and its statistical significance has increased. This is likely to reflect changing practice in schools as the new link with pay makes performance reviews a higher priority for all concerned.

An important part of the policy of objective setting and performance appraisal in schools has been to link classroom performance to the wider performance objectives of the school. The reforms introduced in 2000 took a step in this direction as classroom objectives were meant to be related to the wider goals of the school as set out in its School Development Plan, sometimes also called the School Improvement Plan. All schools are obliged to have such a plan. When Marsden and Belfield (2007) studied the evolution of the 2000 scheme, they found evidence that an increasing minority of schools by 2004 had joined up the different levels of objective setting. If organisations are to have clear goals to guide leadership decisions and which inform the performance reviews of individual employees, then they have

a choice between ‘top-down’ management in which organisational objectives are ‘cascaded down’ to the objectives of individual employees, and a more collegial approach in which organisational objectives inform individual objectives, but employees are involved in establishing organisational priorities. Thus, an important question about the nature of objective setting in schools concerns how far leaders involve classroom teachers in formulating and establishing school priorities, in this case, those of the School Development Plan (SDP) (see for example Pearce et al. 2007, and Nicolaides et al., 2014). In the survey, teachers were asked a number of questions about leadership in their school, and this was the one that was most objective and most specifically related to performance management in schools. The middle panel of Table 5 reports teachers’ views on how well their school leaders involve them in developing the priorities of the School Development Plan. Greater teacher involvement is associated with more favourable views of PRP, and with fewer teachers finding it divisive.

The final panel of Table 5 explores the relationship between systematic appraisals and teacher involvement in the school development plan. Although positive, the relationship is weak and not statistically significant. One reason for this may be that the questions selected for systematic appraisal relate closely to the official guidance, and say little about whether the process is mechanistic or supportive. Remember that the appraisal fairness questions were omitted from this analysis to avoid potential halo effects from views of PRP. Thus it is quite possible that schools with top-down management do appraisal systematically, but do not involve teachers in the overall objectives of the school.

Table 5: Relationship between how teachers view PRP, and systematic appraisal procedures and participatory management in schools[12]

	PRP Motivates 2014				PRP Motivates 2015			
Systematic appraisal procedures	Less positive	More positive	Total %	Sample N	Less positive	More positive	Total %	Sample N
Better than average	47.6	52.4	100	459	51.5	48.5	100	489
Below average	43.0	57.0	100	318	58.3	41.7	100	248
prob Ch2			0.269	-			0.134	+
	PRP divisive 2014				**PRP divisive 2015**			
Systematic appraisal procedures	More divisive	Less divisive	Total		More divisive	Less divisive	Total	
Better than average	49.7	50.3	100	463	41.7	58.3	100	488
Below average	47.2	52.8	100	319	56.2	43.9	100	251
prob Ch2			0.547	-			0.001	****

	PRP Motivates 2014				PRP Motivates 2015			
Leaders involve teachers in SDP	Less positive	More positive	Total %	Sample N	Less positive	More positive	Total %	Sample N
Leaders involve	39.1	60.9	100	155	48.0	52.0	100	170
Leaders don't involve/ not sure	46.9	53.1	100	622	55.9	44.1	100	567
			0.139	+			0.119	+
	PRP divisive 2014				**PRP divisive 2015**			
Leaders involve teachers in SDP	More divisive	Less divisive	Total		More divisive	Less divisive	Total	
Leaders involve	37.9	62.1	100	155	29.7	70.3	100	169
Leaders don't involve/ not sure	51.3	48.8	100	627	52.5	47.5	100	570
prob Ch2			0.013	***			0.000	****

	Leaders involve teachers in developing priorities of the SDP							
	2014				2015			
Systematic appraisal procedures	Leaders don't involve/ not sure	Leaders involve	Total %	Sample N	Leaders don't involve/ not sure	Leaders involve	Total %	Sample N
Better than average	81.8	18.2	100	467	76.3	23.7	100	510
Below average	77.2	22.8	100	325	81.5	18.6	100	282
prob Ch2			0.175	+			0.145	+

Note: panel sample weighted.

7. School management quality and performance pay

One of the major unresolved questions in relation to performance pay and appraisal is their relation to the overall quality of management in an organisation. Often, empirical studies of PRP are confined to a single organisation so it is not easy to explore the effect of variations in management quality. Often too when there are multi-organisational studies, there are significant variations in the design of PRP systems, and there is also a problem of selection bias, that 'better' managed or more successful units are more likely to experiment with new reward practices. The new pay system for teachers avoids many of these problems because the scheme is compulsory for all schools, however they are managed, and local management has some autonomy over how it is implemented. The reports on appraisal and involvement just reviewed provide valuable insights into key processes of school performance management as seen by teachers, but for a more thorough and independent assessment we turn to the national school inspectorate reports conducted by Ofted in England and Estyn in Wales.

These are normally based on evidence provided by the school, and a two-day intensive site visit that involves meetings with school leadership, teachers, students and parents, together with classroom observation. In recent years, school inspectors have been required to enlarge their brief from educational attainments to include the quality of school leadership and management processes with a focus on how they contribute to educational outcomes. As defined in each Ofsted school report, leadership and management is evaluated based on the 'contribution of all the staff with responsibilities (…) to identifying priorities, directing and motivating staff and running the school'. Ofsted's School Inspection Handbook specifies multiple processes and structures across leadership levels. Processes include providing a clear vision; pursuing and modelling high professional standards in teaching and achievement; effective financial management, governance, monitoring and self-evaluation, curriculum development and assessment as well as equality, health and safety regulations (Ofsted, 2015). Inspectors are also required to examine the operation of performance management, and to check that the proportion of awards corresponds to actual achievements by the school. More detail is provided in the appendix. The inspectors' reports include an overall score for management ranging from 'Outstanding' through 'Good', 'Requires improvement' and 'Inadequate' based on the inspectorate's understanding of good educational outcomes and management practices[13]. The reports are placed in the public domain to inform parents and local communities about their schools, and to guide the work of the Education Department in securing high standards for the nation's schools. The inspectors pay special attention to rates of progress as opposed to simple levels of achievement, and in this sense provide a control for some of the factors affective levels of achievement such as deprivation. They are taken very seriously by schools themselves, and a bad report can have very serious consequences for a school and its leadership. The inspectorate knows that poorly conducted inspections may be challenged.

No single index can capture the complexity of a school's life and management, so it is important to ask whether its use by other researchers has led to meaningful results that validate the inspectorate scores as measures of educational and management effectiveness.

On the educational side, Besley at al. (2006) find a relationship between their own indicators of management quality management scores and inspection judgements in a pilot study of schools in England. Hussain (2012) shows that inspectors' 'ratings that make it easier to distinguish between more and less effective schools; and […] they are able to identify poorly performing schools, leading to test score gains', and Allen and Burgess (2012), by comparing marginally failing with marginally passing schools, show that school critical inspection reports lead to improvements in school performance. In addition, the more detailed indicators of management style enabled Currie et al (2007) to develop a typology of models of head teacher leadership.

Table 6: School inspectors' leadership scores compared with teacher evaluations of school leadership and PRP, 2014-2015.[14]

Ofsted Leadership Grade	Leaders involve teachers in SDP	Systematic appraisal procedures	PRP Motivates	PRP divisive	N
1: Outstanding	0.070	-0.048	0.074	0.032	340
2: Good	0.046	0.064	-0.022	-0.043	654
3: Requires Improvement	-0.219	-0.059	0.031	0.047	170
4: Inadequate	-0.537	-0.506	-0.524	0.310	30
Total	0	0	0	0	1194
Anova F	5.33	4.49	4.02	2.63	
p>F	0.001	0.004	0.007	0.049	

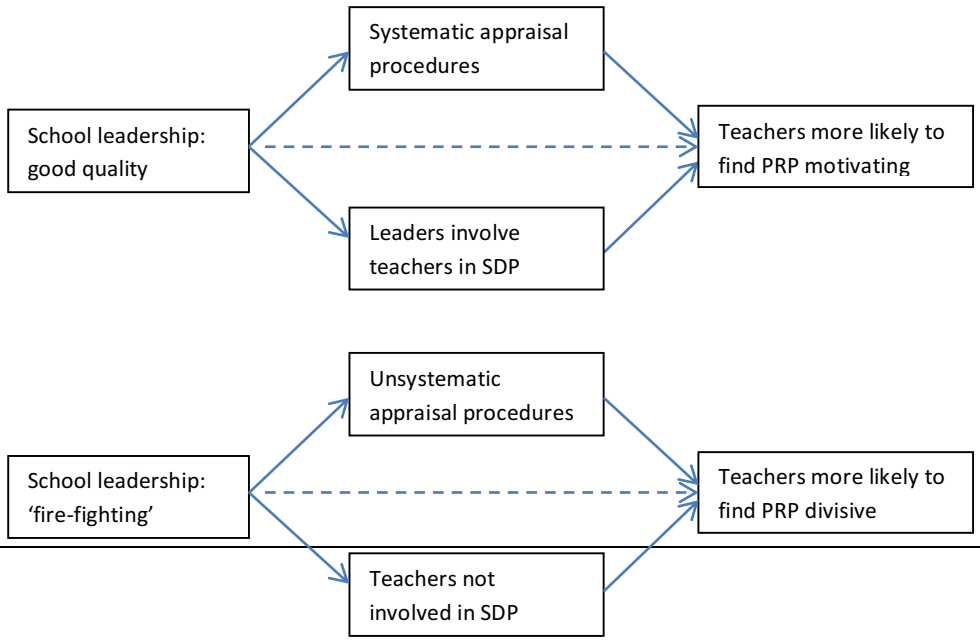
Note: The scores for each variable have been computed such that roughly two-thirds of observations should fall within the range of +1 or -1 about a mean of zero. Information on appraisals and PRP attitudes has been aggregated across questions, see Tables 1 and 2 above, using factor analysis for the full sample over both years. The scores show deviations from the mean score for all Ofsted grades for the panel sample in 2014 and 2015. This was necessary because the factor scores were computed from the full sample including those not in the panel. Involvement in the school development plan was computed from the original five point Likert scale for that variable. The F statistic is computed by ANOVA comparing the variation of scores between Ofsted grades with that within those grades. The probability that the scores by Ofsted grade are statistically significant is shown the bottom row: 0.004 means there is less than 1% (0.4%) probability the overall differences between Ofsted grades are due to chance.

The grades awarded have the merit of providing an overall judgement based on evidence provided by schools and a two-day intensive visit. For the present study, the inspectorate scores have the advantage of providing an external source of information about the quality of management in individual schools that is independent of the survey responses. Moreover, the distribution of grades awarded to schools in the survey correspond quite closely to national inspection results reported by Ofsted (see Appendix 1).

These leadership ratings for each school can be linked to teachers' responses in the survey. Table 6 compares the inspectorate's leadership scores with the teacher replies about whether appraisals are systematic, whether leaders involve teachers in school governance, and whether PRP is experienced as motivating or divisive for 2014-2015. Generally speaking, in schools whose leadership is more highly rated by Ofsted, both teacher involvement and systematic appraisal appear stronger than in those with poor Ofsted ratings. The same is true for PRP: it is seen as more motivating (less demotivating) and less divisive in schools that are better led. The differences are particularly strong compared with schools whose leadership Ofsted judged Inadequate. Thus well-managed schools appear to be more likely to involve teachers more in establishing school-wide priorities, and have systematic appraisal procedures, , and to have less divisive evaluations of PRP. Table 7 seeks to examine all of these factors simultaneously using regression analysis.

The discussion so far has shown that whether teachers find PRP motivating or divisive, the quality of the review procedures, involvement in their school's governance and the external leadership scores are interrelated. Existing research suggests a different relationship between involvement and good procedures on the one hand and whether PRP is seen as motivating or divisive on the other. The reviews by Locke and Latham's (2002) and by Cawley et al (1998) suggest that involving teachers in determining school objectives is likely to motivate them because in this way they take ownership of their work objectives, and it creates a more collegial atmosphere. Thus participative management is more likely to create a basis for shared objectives, and linking pay to their achievement can then be experienced as reinforcing this process. In contrast, according to Greenberg et al (2001), procedural justice theory suggests that poor objective setting and evaluation procedures can undermine teachers' confidence in the fairness of PRP, as they see their chances of reward reduced even when their performance warrants it. Lack of involvement would similarly increase feelings of unfairness. These two models are sketched out in Figure 1.

Figure 1: Possible relationship between school leadership quality and whether teachers find PRP motivating or divisive



In the upper model, good leadership, as evaluated by Ofsted is associated with well-implemented appraisal procedures and involvement of teachers in establishing the priorities of the School Development Plan, and in such an environment, teachers are more likely to find PRP motivating. In the second, weaker leadership, again as evaluated by Ofsted, is associated with less well implemented performance review procedures and less involvement of teachers in the SDP, and in this environment teachers are more likely to find PRP divisive.

These two models are compared with the survey data in Table 7, using a two-level regression which estimates the effect of management quality on involvement and appraisal, and then the effect of the latter two on teachers' views of PRP In this procedure, 'seemingly unrelated regression', the presence of an 'effect' does not necessarily imply causation. Considering the upper panel, on the right hand side, in schools where leaders involve teachers in developing the priorities of the school development plan, which is a central reference point for teachers' performance reviews, teachers are nearly 20% (0.193) more likely to report PRP as motivating than in other schools. Likewise, they are nearly 20% (-0.189) less likely to find PRP divisive. In contrast, the effect of systematic appraisal procedures is not significant for motivation, but it is related to feelings of divisiveness (about 10% more, 0.112). The rest of the upper panel confirms the effects of teachers' opportunities for performance based progression by virtue of their position on the pay spine discussed earlier.

The lower panel shows the link between teacher involvement and appraisal procedures on the one hand, and the quality of school leadership as assessed by the most recent school inspection. Taking schools with an 'Outstanding' rating as the benchmark, the left-hand side shows that the relationship with SDP involvement declines as the Ofsted grade deteriorates. Thus compared with an 'Outstanding' school, teachers in one graded 'Inadequate' are 50% less likely to experience SDP involvement, and consequently will find PRP about 8-9% less motivating than other teachers (-0.50 x 0.17). Turning to appraisals, it appears that both Outstanding and Inadequate schools are less well-equipped with systematic procedures than the others. This may be partly because of the small sample size where a few exceptions may distort the result. It may also reflect a general tightening up of statutory appraisal, and especially feedback procedures in many schools (Table 4 above) made necessary by the link with pay. It may also reflect the focus on statutory aspects of appraisal quality which were used for methodological reasons discussed earlier, and which do not distinguish between the spirit and the letter of appraisal rules.[15]

Use of inspectors' management grades requires some caution in interpreting the results of the present study. As inspections normally operate on the three-yearly cycle, it is likely that adverse reports from early in the cycle have been followed by improvements in school management and in educational results. Equally, recent but unexpectedly negative inspections may have an adverse effect on teacher morale, and colour some attitudes measured in the current survey. However, in the latter case, the effect on teacher morale is more likely to stem from the inspectors' overall educational grading, rather than the narrower grading for leadership.

Table 7: Perceptions of PRP as motivating or divisive by spine point and school management style 2015[16]

	PRP motivates			PRP divisive		
	Coef.	Std. Err.	sig	Coef.	Std. Err.	sig
Leaders involve teachers in SDP	0.173	0.033	****	-0.176	0.028	****
Systematic appraisal procedures	0.030	0.035	-	-0.113	0.030	****
Primary school	-0.041	0.077	-	0.055	0.065	-
Part-time	-0.137	0.082	*	0.027	0.070	-
Female	0.092	0.070	+	-0.002	0.060	-
Spine point						
Spine pt M5	-1.104	0.325	****	0.156	0.276	-
Spine pt M6	-0.419	0.162	****	0.134	0.138	-
Spine pt U1	-0.166	0.172	-	0.068	0.147	-
Spine pt U2	-0.048	0.168	-	0.174	0.143	-
Spine pt U3	-0.310	0.094	****	0.221	0.080	****
Spine pt Ldr1-5	-0.013	0.219	-	0.021	0.186	-
Movement up a spine Point 2014-2015	0.239	0.122	**	-0.206	0.103	**
Spine point in 2015 & receipt of award						
Rise to M5	0.943	0.374	***	-0.100	0.318	-
Rise to M6	0.393	0.227	*	0.136	0.193	-
Rise to U1	0.618	0.230	****	-0.353	0.195	*
Rise to U2	-0.048	0.197	-	-0.116	0.167	-
Rise to U3	0.023	0.116	-	-0.067	0.099	-
Rise within L1-5	0.402	0.288	+	-0.339	0.245	+
Constant	-0.383	0.119	****	0.374	0.101	****
N		781			780	
R2		0.1002	****		0.0992	****

	Leaders involve teachers in SDP			Systematic appraisal procedures		
Ofsted leadership score	Coef.	Std. Err.	sig	Coef.	Std. Err.	sig
Grade 2	-0.019	0.076	-	0.271	0.072	****
Grade 3	-0.232	0.118	**	0.234	0.112	**
Grade 4	-0.501	0.256	*	-0.204	0.244	-
Primary school	0.281	0.081	****	-0.025	0.077	-
Constant	2.499	0.053	****	0.184	0.051	****
R2		0.0267	****		0.0215	****
N		781			781	

Notes: SURE: Seemingly unrelated regression. Missing values for leaders involve in SDP and for appraisal procedures for 2015 were dealt with by using responses for 2014 and taking the average across both years for all respondents. Panel sample, unweighted. Ofsted Grade 1 is benchmark grade.

8. Is stronger perceived motivation of PRP related to higher performance?

How employees feel about their pay systems, whether they find them motivating and fair, is an important aspect of their job satisfaction, but it does not necessarily affect their job performance. This section explores whether there is a link between teachers' prior views about whether they find PRP motivating (reported in the 2014 survey), and their subsequent achievement of the objectives set in their performance reviews (reported in the 2015 survey). As discussed earlier, these reviews are designed to cover a wide range of aspects of classroom performance, including pupil progress, as established in the objective setting meeting. The criteria are judgemental, and so schools may vary in how specifically objectives are defined, and how strictly they are assessed. However, such variation is limited by the obligation on school inspectors to examine the overall relationship between school performance and the percentage of positive reviews, and to challenge significant divergences. This is an input into their grading of school leadership performance. Additionally, many teachers are at the top of their respective pay scales so that they can only progress by promotion, through the Threshold onto the Upper Pay Scale, or onto the Leadership Scale. Although these are all part of the performance management system in schools, they involve a more complex evaluation than the annual cycle of objective setting and performance review.

If the system is working as designed, then teachers who meet the objectives of their performance reviews should receive a recommendation on pay, which if validated by the leadership and governing body, should lead to pay progression. Thus, with the exception of those at the top of their respective pay scales, one may use the receipt of pay progression as an indicator that their school recognises that they have met their performance objectives. This is examined in Table 8, which considers perceived motivation and spine position in the previous year, 2014, and subsequent achievement of objectives as reported in 2015. We have therefore a measure of whether teachers who said they found PRP motivating at the start of the cycle, went on to achieve their objectives by the end, a year later.

The results in the upper panel show that teachers who found PRP motivating in 2014 were about 45% ((1.444 – 1) x 100) more likely to have met their objectives than other teachers who had received either a negative recommendation, or no recommendation at all. To put this into perspective, roughly a fifth of the panel teachers had responded positively to the motivational questions (Table 1 above), the others being either uncertain or negative. In addition, only about a third of the panel sample reported having received a recommendation, and of these roughly half had progression, representing about a sixth of the overall panel sample. [17]

Being on the scale maximum (M6 for the Main and U3 for the Upper Scale) in 2014 has the expected effect: pay recommendations are less likely (coefficients below one) whether or not for progression. Variation between schools in the strictness of reviews and progression is captured potentially in the link between higher rates of progression in schools that inspectors awarded lower inspection grades for their leadership.

Table 8. Comparison of perceived motivation in 2014 with appraised performance in 2015: *benchmark: teachers reporting recommendation on pay*[18]

Panel A: without interactions

	Received a recommendation on pay in 2015					
	No pay progression awarded in 2015 (objectives presumed not met)			Pay progression awarded in 2015 (objectives presumed achieved)		
	Coef. RRR	Std. Err.		Coef. RRR	Std. Err.	
PRP motivates 2014	1.030	0.132	-	1.444	0.173	****
Primary school	1.345	0.384	-	1.272	0.356	-
Part-time	1.013	0.299	-	0.595	0.199	+
Female	0.723	0.194	-	0.863	0.227	-
On main scale max: 2014	0.328	0.163	**	0.771	0.244	-
On upper scale max: 2014	0.682	0.153	*	0.015	0.015	****
Ofsted leadership 2	1.357	0.340	-	1.642	0.451	*
Ofsted leadership 3	1.064	0.425	-	2.249	0.862	**
Ofsted leadership 4	0.393	0.418	-	0.306	0.347	-
Constant	0.315	0.079	****	0.277	0.073	****
Number of obs	777					
Pseudo R2	0.0892					

Table 8: Panel B: including interactions between motivation and being on a scale maximum

	(objectives presumed not met)			(objectives presumed achieved)		
	Coef. RRR	Std. Err.		Coef. RRR	Std. Err.	
PRP motivates 2014	0.890	0.168	-	1.251	0.162	*
Primary school	1.349	0.384	-	1.279	0.363	-
Part-time	0.985	0.289	-	0.552	0.190	*
Female	0.715	0.190	-	0.856	0.229	-
On main scale max: 2014	0.344	0.171	**	0.634	0.236	-
On upper scale max: 2014	0.682	0.152	*	0.005	0.005	****
M6 in 2014 & PRP motivates	1.828	0.756	+	3.019	1.249	****
U3 in 2014 & PRP motivates	1.353	0.325	-	0.149	0.030	****
Ofsted leadership 2	1.328	0.332	-	1.584	0.438	*
Ofsted leadership 3	1.090	0.440	-	2.412	0.951	**
Ofsted leadership 4	0.392	0.413	-	0.306	0.344	-
Constant	0.322	0.082	****	0.291	0.076	****
Number of obs	777					
Pseudo R2	0.0979					

Notes: Coefficients in columns 2 and 5 compare those who reported receiving a recommendation on pay with those who did not, and are based on a multinomial logit regression with 595 observations and a pseudo-R2 of 0.0565. Coefficients should be read horizontally, and compare with the benchmark case of no progression and no recommendation. The dependent variable comprised three categories: No recommendation on pay in 2015, Recommendation on pay 2015 a) without pay progression, and b) with pay progression. It is presumed that most negative recommendations result in an absence of pay progression, and most positive ones, in progression.

The lower panel of Table 8 separates the different effects of motivation for along-scale progression from between-scale progression at the Threshold. This is done by introducing an interaction between perceived motivation and whether teachers were at the top of the Main or the Upper Scales at the start of the cycle, in 2014. Teachers who said PRP was motivating at the outset were about three times more likely to have achieved threshold progression at the end of the cycle. Progression at M6 entails movement to U1, in other words, success in passing the Threshold. However, separating out the Threshold in the model almost halves the effect of motivation on progression in general shown in the upper panel from 45% (1.44) to 25% (1.251), still important, but potentially weaker than that of Threshold progression.

Thus a tentative conclusion is that strong motivation is associated with good performance among the teachers who are eligible for PRP, and that the effect of progression at the Threshold is stronger than for progression along the current pay scales. Nevertheless, this is subject to some 'health warnings'. First, PRP may be acting as a 'lightning rod' for other problems, such as workload, which could cause teachers to be cynical about PRP and find their objectives hard to achieve. Second, the achievement of objectives is inferred from the combination of a pay recommendation and pay progression. Although the Education Department's guidance to schools is that progression should follow performance, it is possible that some schools may delay progression because of financial pressures. Many respondents were of this view (see Table 2, last question), and this has also been reflected in casework by the teachers' unions. Third, it is also possible that some schools are more demanding than others in the award of good appraisal ratings and pay progression. An indication of this may be seen in Table 8 in the coefficients on the 'Ofsted leadership grades' in the right-hand panel. These suggest that compared with grade 1 'Excellent' schools, those with grade 2, and even more so, those with grade 3, are more likely to award pay progression. Note that the remit of inspectors includes the rigour with which performance management is operated. Finally, at just under 800, the sample is quite small.

9. Conclusions

The study shows that both the overall design and local management operation of performance pay are important. As concerns design, one of the most striking findings is the differentiation of teachers' perceptions of PRP according to where they stand on the teachers' pay spines: whether or not this gives them scope to benefit from performance related progression, whether or not they are at the top of their pay scale, and if they were at other points, whether they feel they have lost out as a result of loss of automatic increments.. Few organisations introduce performance pay in a greenfield location. It is much more common to do so in a way that involves movement from an established reward system for incumbent employees, and management teams. How the opportunities to benefit from the new incentives are distributed has been seen to have a clear effect on perceptions of motivation and divisiveness. There is also some evidence, relating to design, that teachers feel more at ease with pay linked to promotion than with annual performance objectives, as reflected in Threshold as compared with performance related progression.

The more important question relates to operation and management quality. In large-scale organisations with many local units, top management has to work with local management teams with differing levels of expertise and effectiveness. In so far as the school inspectorate provides a good measure of these, it is clear that employees in better managed units are more likely to find performance related pay motivating, and less so where management is more engaged in fire-fighting. Moreover, the study suggests that involvement of employees and to a lesser extent, having systematic appraisals, are important channels through which more effective management may take effect. Finally, do employee attitudes to PRP matter, or are they like the weather, a subject of eternal complaint, but something people learn to live with? The latter section of the paper suggests that teachers who appeared more motivated in 2014 when the scheme was introduced were more likely to achieve their objectives than those who were not. The same analysis suggests that school leaderships that were deemed by inspectors to be less effective were more likely to award pay progression, possibly reflecting retention and other pressures on them.

As with all empirical studies, there are limitations, notably in this case in terms of the response. Teachers' participation in this kind of survey is a demand on scarce free time, and although discussion with the teachers' unions suggests that the response is relatively good for an unpaid online survey, it nevertheless is a limitation in terms of potential bias and statistical significance. Finally, although teaching, like many highly educated professions, provides scope for high levels of job discretion, and intrinsic motivation, it is likely to attract people who value these characteristics, and this may colour their beliefs about and attitudes towards performance pay.

10. Appendix 1: Explanation of the Ofsted grades for management in schools

Ofsted inspectors visit schools regularly about every three years, and more frequently for schools in difficulty. Normally, there is a team of inspectors with one lead inspector, and visits last about two days, involving a review of documentary evidence on school policies, including the *School Development/Improvement Plan* and *Self-Evaluation* documents, outcomes, and pupil performance, interviews with selected role holders notably in senior and middle management, classroom observations, and feedback from pupils and parents (Ofsted 2015). Inspectors have a specific list of evidence items to review, which is set out in detail in the Ofsted School Inspection Handbook.[19] They are required to assess overall effectiveness in terms of the quality of education provided of the schools they visit. This covers a number of domains including the quality of leadership and management; the behaviour and safety of pupils; the quality of teaching; achievement of pupils; and the effectiveness of early years and of sixth form provision. For each of these domains, there is a set of descriptors for the four grades used: outstanding, good, requires improvement, and inadequate. There is a strong focus on outcomes and where processes are evaluated it is usually in relation to targeted outcomes.

Inspections focus primarily on quality of education for the school's pupils, but in recent years, Ofsted has been given the brief to examine the quality of leadership and management in so far as they contribute to educational quality. This is reflected in the opening sentences of the section on management: 'In reaching their judgement on leadership and management, inspectors should consider the school's use of performance management and effectiveness of strategies for improving teaching, including the extent to which the school takes account of the 'Teachers' Standards'' (Ofsted 2015 January no. 120101: §159).

There then follows a list of items through which effective leadership is demonstrated, and which inspectors should investigate:

- effectiveness of procedures for monitoring the quality of teaching and learning, and tackling underperformance;
- the strength of the link between performance management, appraisal and salary progression;
- how well leadership groups manage staff performance and use the staff budget to differentiate appropriately between high and low performers;
- coherence and effectiveness of professional development, and opportunities provided for promotion, and its link to identified needs of staff;
- accuracy with which best practice is identified, modelled and shared;
- whether less than good performance is rigorously managed and appropriate training and support are provided; and whether good performance is recognised through the performance management process;
- whether anonymised information about performance management, appraisal and salary progression corresponds to numbers of teachers meeting their performance management objectives for the previous three years.

Initially, the grade descriptors for 'outstanding' and 'good' schools include a section on the quality of performance management, and its contribution to improvements in teaching and learning, and in professional development for all staff. However, in the revised version of Sept 2015, the references to performance management contained the same principles but were less prescriptive (§138).

The distribution of Ofsted grades as in 2014 shows that the modal grade is 'Good', with similar percentages of schools classed as 'Outstanding' and 'Requires improvement'. The share of schools classed 'Inadequate' is very small. The distribution of schools by Ofsted grade in the current survey is very close to the national picture (Appendix Table 1).

Appendix Table 1: Overall effectiveness of open maintained schools and academies at their most recent inspection as at 31 August 2014

School type	No of schools inspected	Outstanding	Good	Requires improvement	Inadequate
Primary schools (total)	16,266	17	64	16	2
CEP Survey 2014		*16*	*68*	*16*	*0*
Secondary schools(total)	3,150	21	49	23	6
CEP Survey 2014		*26*	*50*	*20*	*4*

Source: Ofsted (2014: 36) and CEP survey.

11. Appendix 2: Survey method and potential sample selection bias

Teachers' views were obtained by means of a random sample survey sent by the teachers' unions to about 10% of their membership with the addresses drawn from their membership lists. Unionisation among school teachers is very high, at roughly 90%[20], so the effect of any bias should be small. The survey was administered in the spring months of each year. Participation was voluntary, and respondents were asked to compose their personal confidential id code according to a fixed method which could be replicated for future replies, and so used to link replies without providing personal details. Those who wished to take part in the follow-up were also asked if they would provide their contact email. Respondents were also asked to provide the name and postcode of their schools so that their replies could be linked to school level data from other sources, such as those of school inspector reports and the DfE. To boost the panel element, the unions were asked, if possible, to use the same sample of contact emails from year to year, combined with a refreshment sample. The total usable response in 2014 was just under 3,000 (4,600 in total), and in 2015, 1,000 (1,300 in total), and about 790 records could be linked between the two years for the panel.

Comparison of certain demographic characteristics of respondents and details of schools between the 2014 survey and the School Workforce Census[21] showed good coverage of the teacher population and school types (see below and Marsden, 2015). The most notable

divergence was that the response was notably higher among more senior teachers on the higher points of the school teachers' salary scale. For this reason, many of the analyses in this paper have been weighted to give a truer reflection of the overall population of teachers.

Comparison of sample response by salary spine point with School Workforce Census 2013

Salary Spine point	Full sample	Panel sample	School Workforce Census
	2014	2014/2015	2013
	%	%	%
UQT	0.0	0.0	n.a
M1-2	5.5	5.2	16.1
M3-4	8.2	8.2	13.4
M5	4.4	4.4	6.1
M6	**13.3**	**9.9**	**12.6**
U1	12.5	13.2	13.3
U2	12.3	12.2	9.8
U3	**39.3**	**43.2**	**25.8**
L1-3	4.6	3.6	2.9
Total	100	100	100
Total	2648	745	

Note: responses giving sufficient detail on pay to enable matching.

Classroom teachers: demographic characteristics

Characteristic	2014 Response composition %	School work-force census 2012 (%)	Characteristic	2014 Response composition %	School work-force census 2012 (%)
Gender			Age	%	
Female	71.1	75	Under 25 years	2.4	6
Male	28.9	25	25-29	10.0	18
Total %	100	100	30-34	12.7	17
N	2876		35-39	13.3	14
			40-44	15.7	13
Employment status			45-49	16.0	11
Full-time	81.8	76	50-54	15.1	10
Part-time	18.2	24	55-59	11.4	9
			60 and above	3.4	2
Total %	100	100	Total %	100	100
N	2875		N (complete replies)	2883	

Sample selection bias is a concern with all voluntary response surveys, and in the present survey, that the degree of discontent among teachers is exaggerated as the most discontented my reply in greater numbers. This cannot be tested directly, but limited comparison is possible with the teacher respondents in the 2011 Workplace Employment Relations Survey who replied to a number questions about their schools covering views of relations with their managers, including trust, commitment, and shared values. The same questions were included

in the CEP survey, in part, to enable benchmarking. The results are shown below for the first wave of the survey, and include comparisons with an earlier CEP study of teachers in 2000 (for fuller details see Marsden, 2015, Table 5). The results in the CEP surveys are quite close to each other and to those of WERS, albeit showing slightly stronger negative attitudes compared with WERS. Moreover, in discussions with the teachers' unions, it was suggested that teachers' attitudes could have been more negative, especially in 2014 owing to discontent over increased workloads.

Thus, comparison with WERS results for teachers suggests that there may be some selection bias in terms of more discontented teachers took the time to complete the questionnaire, but it is at worst moderate, and it is also likely that a share of the increased negative view compared with 2011 is the result of uncertainty surrounding the new pay system and the problems of teachers' workloads, the latter being recognised by the employers.

Teachers' attitudes towards their schools: a comparison with other sources.

	Disagree	Neutral	Agree
I am proud to tell people who I work for			
Teachers 2014 (CEP)	12.3	21.6	66.1
Teachers 2000 (CEP)	10.7	21.5	67.8
WERS 2011 teachers	5.3	14.1	80.6
WERS 2011 all employees (excl teachers)	9.3	23.2	67.6
I share many of the values of my organisation			
Teachers 2014 (CEP)	11.4	14.1	74.6
WERS 2011 teachers	3.9	8.9	87.2
WERS 2011 all employees (excl teachers)	7.9	27.7	64.4
I feel loyal to my organisation			
Teachers 2014 (CEP)	12.9	15.3	71.8
WERS 2011 teachers	4.8	8.1	87.0
WERS 2011 all employees (excl teachers)	7.9	17.5	74.6
Using my own initiative I carry out tasks that are not required as part of my job			
Teachers 2014* (CEP)	3.5	28.0	68.5
WERS 2011 teachers	3.8	12.4	83.9
WERS 2011 all employees (excl teachers)	9.1	20.4	70.5

Notes: * Teachers 2014 classified responses on initiative as 'never', 'sometimes' 'quite often' and 'very often'. 'Sometimes' was classified as 'neutral' and quite and very often as 'agree'.
Source: Marsden 2015, Table 5.

12. References

Aaronson, Daniel; Barrow, Lisa, and Sander, William (2007) Teachers and Student Achievement in the Chicago Public High Schools. Journal of Labor Economics, 25: 1, January, pp. 95-135.

Allen, Rebecca, and Burgess, Simon (2012) How should we treat under-performing schools? A regression discontinuity analysis of school inspections in England. *CMPO Working Paper No. 12/287*, Centre for Market and Public Organisation, University of Bristol.

Atkinson, Adele; Burgess, Simon;, Croxson, Bronwyn;,Gregg, Paul; Propper, Carol; Slater, Helen and Wilson, Deborah (2009) Evaluating the impact of performance-related pay for teachers in England. Labour Economics, 16: 3, June pp. 251-261

Bakker, Arnold B.; Hakanen, Jari J.; Demerouti, Evangelia; Xanthopoulou, Despoina (2007) Job resources boost work engagement, particularly when job demands are high. Journal of Educational Psychology, 99: 2, May, pp. 274-284.

Besley, Timothy; Hussain, Iftikhar, and Machin, Stephen (2006) Do public sector CEOs make a difference? Results from a pilot survey of headteachers. *Working Paper, Centre for Economic Performance*, London School of Economics, London.

Bloom, Nicholas; Propper, Carol; Seiler, Stephan, and Van Reenen, John (2010) The Impact of Competition on Management Quality: Evidence from Public Hospitals CEP Discussion Paper No. 983, May, London School of Economics, London.

Bloom, Nick; Sadun, Raffaella and Van Reenen, John (2007) Americans Do I.T. Better: US Multinationals and the Productivity Miracle. CEP Discussion Paper No 788, April, Centre for Economic Performance, London School of Economics.

Brown, Wilfred (1962) *Piecework abandoned: the effect of wage incentive systems on managerial authority*. Heinemann, London.

Cawley, Brian; Keeping, Lisa; and Levy, Paul (1998) Participation in the Performance Appraisal Process and Employee Reactions: A Meta-Analytic Review of Field Investigations; *Journal of Applied Psychology*, 83: 4, August, pp. 615-633.

Cropanzano, Russell; Bowen, David E. and Gilliland, Stephen W. (2007) The Management of Organizational Justice. Academy of Management Perspectives, 21: 4, Nov., pp. 34-48.

Currie, Graeme; Lockett, Andy, and Suhomlinova, Olga (2009) Leadership and institutional change in the public sector: The case of secondary schools in England. *The Leadership Quarterly*, 20: 5, Oct, pp. 664-679.

DfE (2012) *The Education (School Teachers' Appraisal) (England) Regulations 2012*, Statutory Instruments, 115, 2012. Department for Education, London.

DfE (2013) *Departmental advice: Reviewing and revising your school's approach to teachers' pay*. Dept for Education, London.

DfE (2014a) *Implementing your school's approach to pay: Departmental advice for maintained schools and local authorities*. Department for Education, London.

DfE (2014b) *School teachers' pay and conditions document 2014 and guidance on school teachers' pay and conditions*, Department for Education, London.

DfE (2015) *School teachers' pay and conditions document 2015 and guidance on school teachers' pay and conditions*, Department for Education, London.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan (2012) Incentives Work: Getting Teachers to Come to School. *American Economic Review*: 102: 4, June, pp. 1241-78.

Erdogan, B, Kraimer M.L., and Liden R.C. (2001) Procedural Justice as a Two-Dimensional Construct: An Examination in the Performance Appraisal Context. *The Journal of Applied Behavioral Science,* June, 37: 2, pp. 205-222.

Figlio, David N. and Kenny, Lawrence W. (2007) Individual teacher incentives and student performance. *Journal of Public Economics*, 91: 5–6, June, pp. 901–914.

Gabris G. T, and Ihrke D. M. (2000) Improving employee acceptance towards performance appraisal and merit pay systems: the role of leadership credibility. *Review of Public Personnel Administration*, Columbia.

Gagné, Marylene and Forest, Jacques (2008) The Study of Compensation Systems Through the Lens of Self- Determination Theory: Reconciling 35 Years of Debate. *Canadian Psychology*, 49: 3, pp. 225–232.

Greenberg, Jerald, and Cropanzano, Russell, eds. (1999) *Advances in Organizational Justice*. Stanford University Press, Stanford CA.

Gruman, Jamie A. and Saks, Alan M. (2011) Performance management and employee engagement. *Human Resource Management Review*, 21: 2, June pp.123-136.

Hanushek, Eric A, and Rivkin, Steven G. (2010) Generalizations about Using Value-Added Measures of Teacher Quality. *American Economic Review*, 100: 2, May, pp. 267-271.

Hussain, Iftikhar (2012) Subjective Performance Evaluation in the Public Sector: Evidence from School Inspections. *CEE Discussion Paper 135*, Centre for the Economics of Education, London School of Economics, London.

Kahneman, Daniel, and Tversky, Amos (2000) *Choices, values and frames*. Cambridge University Press, Cambridge.

Latham, Gary P, and Pinder, Craig, C. (2005) Work motivation theory and research at the dawn of the twenty-first century; *Annual Review of Psychology*, 56, pp. 485-516.

Lavy, Victor (2009) Performance Pay and Teachers' Effort, Productivity, and Grading Ethics. *American Economic Review*, 99: 5, Dec. pp. 1979-2011.

Lawler E. E. III. (1971) *Pay and organisational effectiveness: a psychological view*. McGraw-Hill, New York.

Lazear, Edward P. (1996) Performance pay and productivity. NBER Working Paper 5672.

Lazear, Edward P. (2007) Personnel economics of managers, 2nd edn. Wiley, New York.

Levy, Paul E., and Williams, Jane R. (2004) The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management*, 30: 6, pp. 881-905.

Locke, Edwin A. and Latham, Gary P. (2002) Building a practically useful theory of goal setting and task motivation: a 35-year odyssey; *American psychologist*, 57: 9, Sept., pp.705-717.

Marsden D. W., and French S. (1998) *What a performance: performance related pay in the public services*. Centre for Economic Performance Special Report, London School of Economics, London, available online from http://cep.lse.ac.uk/pubs/download/special/performance.pdf

Marsden David W. (2015) Teachers and Performance Pay in 2014: First Results of a Survey. *Discussion Paper 1332, Centre for Economic Performance, LSE*. Available online at: http://cep.lse.ac.uk/pubs/download/dp1332.pdf

Marsden, David and Belfield, Richard. (2007) Pay for performance where output is hard to measure: the case of performance pay for school teachers. *Advances in Industrial and Labor Relations*, vol 15, pp 1-37. Available also at http://cep.lse.ac.uk/pubs/download/dp0747.pdf

Muralidharan, Karthik and Sundararaman, Venkatesh (2011) Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 119: 1, February, pp. 39-77.

Murnane, Richard J. and Papay, John P. (2010) Teachers' Views on No Child Left Behind: Support for the Principles, Concerns about the Practices. *The Journal of Economic Perspectives*, 24: 3 Summer, pp. 151-166.

Murphy, Richard (2013) *Testing Teachers: What works best for teacher evaluation and appraisal*, The Sutton Trust, London.

Nicolaides, Vias C.; LaPort, Kate A.;Chen, Tiffani R., Tomassetti, Alan J.; and Weis, Eric J. (2014) The shared leadership of teams: A meta-analysis of proximal, distal, and moderating relationships, *Leadership Quarterly*, 25, pp. 923-942.

OECD (2005) Teachers matter: attracting, developing and retaining effective teachers. Organisation for Economic Cooperation and Development, Paris.

Ofsted (2014) *The report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2013/14*. The Office for Standards in Education, Children's Services and Skills (Ofsted), Manchester. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/384707/Ofsted_Annual_Report_201314_Schools.pdf

Ofsted (2015) *School Inspection Handbook*: *Handbook for inspecting schools in England under section 5 of the Education Act 2005 (as amended by the Education Act 2011)*, The Office for Standards in Education, Children's Services and Skills (Ofsted), Manchester. https://www.gov.uk/government/publications/school-inspection-handbook

January 2015 edition: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/457199/School_inspection_handbook.pdf

Pearce, Craig L.; Conger, Jay A.; and Locke, Edwin A. (2007) Shared leadership theory. *Leadership Quarterly* 18, pp. 281–288.

Podgursky, Michael J. and Springer, Matthew G. (2007) Teacher Performance Pay: A Review. *Journal of Policy Analysis and Management*,26: 4, Autumn, pp.

Rivkin, Steven G; Hanushek, Eric A., and Kain, John F. (2005) Teachers, Schools, and Academic Achievement. *Econometrica*, 73: 2 March, pp. 417-458.

Rockoff, J.E., and Speroni, C. (2010) Subjective and Objective Evaluations of Teacher Effectiveness. *American Economic Review*, 100: 2, May, 261-66

Schmiede, Rudi, and Schudlich, Edwin (1976) *Die Entwicklung der Leistungsentlohnung in Deutschland: eine historisch-theoretische Untersuchung zum Verhältnis von Lohn und Leistung unter kapitalistischen Produktionsbedingungen*. Campus Verlag, Frankfurt.

School Teachers Pay Review Body (STRB) (2012) *Twenty-first report* CM8487, The Stationery Office, London.

Shapiro, Debra, L, and Kirkman, Bradley L. (2001) Anticipatory injustice: the consequences of expecting injustice in the workplace. Chapter 5 in Greenberg, Jerald, and Cropanzano, Russell, eds.(2001) *Advances in Organizational Justice*, Stanford University Press, Standford, CA. pp. 152-178.

Slater, Helen; Davies, Neil M., and Burgess, Simon (2012) Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England. *Oxford Bulletin of Economics and Statistics*, 74: 5, October, pp. 629-645,

Staiger, Douglas O., and Jonah E. Rockoff (2010) Searching for Effective Teachers with Imperfect Information. *Journal of Economic Perspectives*. 24: 3, pp. 97-118.

Taylor, Eric S. and Tyler, John H. (2011) The Effect of Evaluation on Performance: Evidence from Longitudinal Student Achievement Data of Mid-career Teachers. *NBER Working Paper No. 16877*, National Bureau of Economic Research, Cambridge, Mass.

Taylor, M. Susan; Tracy, Kay B.; Renard, Monika K; Harrison, J. Kline; and Carroll, Stephen J. (1995) Due process in performance appraisal: a quasi-experiment in procedural justice. *Administrative Science Quarterly*, 40: 3, Sept., pp. 495-523.

13. Endnotes

[1] Revised version of document 'Teachers PRP 2014-2015 COPE 20170111.docx'.

[2] However, this view was not universally held: many in the teachers' and head teachers' associations argued that the observed high pass rates were due to pre-selection by schools of their candidates for advancement.

[3] A full account of the survey methods and checks on response balance can be found in Marsden (2015).

[4] See School Workforce Census for 2013, Tables 7a and 7b, and method appendix to this paper.

[5] Cronbach's Alpha, measure of internal consistency for an index based on several separate questions, for the PRP motivates index was 0.85 and for the divisiveness one, 0.65 when computed on the whole sample.

[6] Strictly speaking, they were not significantly different from the benchmark category of junior teachers on M1 and M2. The views of this category were close to the average for all teachers, that is negative overall, but not more so than teachers on the intervening spine points.

[7] Program: reg ppgood4_1415 sal_rise lag_spine_max sal_lag_m6 sal_lag_u3 part_time female sk_primary if year==2015 & panel==1 , robust; reg ppdiv4_1415 sal_rise lag_spine_max sal_lag_m6 sal_lag_u3 part_time female sk_primary if year==2015 & panel==1 , robust.

[8] Program: reg ppgood4_1415 m3_4-l1_5 rise_m3_4 rise_m5 rise_m6 rise_u1 rise_u2 rise_u3 rise_l1_5 part_time female sk_primary [pweight=panel_wt2] if year==2015 & panel==1 , robust

[9] Progams: tab2 year objspec if panel==1, row chi2, for variables: objspec expmonit objpupprog oppperf objsdp exp_payknow ppfedbak ppreson ppprevious ppclassobs pp_payrec

[10] An analysis of who had a performance review in 2014 and 2015 showed that most of those who had not had one in 2014 did have one in 2015, suggesting that in such cases, the system was less well embedded in their schools.

[11] Cronbach's Alpha, for 'good procedures', was 0.95 for the whole sample, indicating a high degree of consistency among the questions selected for the index.

[12] programs used: bysort year: tab2 obj_proc_1415_bin ppgood1_1415_bin , row chi2; bysort year: tab2 obj_proc_1415_bin ppdiv1_1415_bin , row chi2; bysort year: tab2 ldrs_sdp_good ppgood1_1415_bin, row chi2; bysort year: tab2 ldrs_sdp_good ppdiv1_1415_bin, row chi2; bysort year: tab2 obj_proc_1415_bin ldrs_sdp_good , row chi2

[13] See the School Inspection Handbook, (Ofsted, 2015) and http://www.clerktogovernors.co.uk/ofsted/ofsted-grade-descriptors-quality-of-leadership-in-and-management-of-the-school/

[14] Programs: use "C:\Users\MARSDEND\Dropbox\DATA_TEMP\Teachers_prp_anon_2014_15.dta", clear; tabstat ppgood1_1415 if panel==1 , statistics(mean semean count) by(ofsted_leadership_14_15); oneway ppgood1_1415 ofsted_leadership_14_15 if panel==1; variables: ppgood1_1415, ppdiv1_1415, ldrs_sdp, obj_proc_1415, ofsted_leadership_14_15.

[15] The r2 statistic measures both the percentage of the variation in the leadership involvement and the systematic appraisal variables accounted for by the Ofsted management quality grades and the amount of 'noise' in its measurement. Because the number of teacher respondents in each school was very small, often only one, the amount of 'noise' will have been great. Such noise arises because there is variation in teacher perceptions within every school, and it is inevitable that respondents' views may not always be typical of all teachers in their schools. The value of the coefficient on management grade conveys information about the strength and size of its relationship with teacher involvement and appraisal quality, whereas the low r2 statistic indicates the amount of noise in that relationship.

[16] Programs: sureg (ppgood4_1415 = ldrs_sdp_av obj_proc_new_av sk_primary part_time female uqt-u2 sal_rise sal_uqt-sal_u2) (obj_proc_new_av = ofsted_ldr_2 ofsted_ldr_3 ofsted_ldr_4 sk_primary year) (ldrs_sdp_av = ofsted_ldr_2 ofsted_ldr_3 ofsted_ldr_4 sk_primary)if panel==1 & year==2015; sureg (ppdiv4_1415 = ldrs_sdp_av obj_proc_new_av sk_primary part_time female uqt-u2 sal_rise sal_uqt-sal_u2) (obj_proc_new_av = ofsted_ldr_2 ofsted_ldr_3 ofsted_ldr_4 sk_primary year) (ldrs_sdp_av = ofsted_ldr_2 ofsted_ldr_3 ofsted_ldr_4 sk_primary)if panel==1 & year==2015

[17] The sample was weighted to compensate for uneven response rates by salary spine point – to compensate for over-reporting by those at the top of their pay scales. About 68% of the weighted panel sample in 2015 reported no recommendation on pay, and those who reported a recommendation divided roughly evenly between not getting and getting progression (17% and 14%).

[18] mlogit rise_payrec_r lag_ppgood1_1415 sk_primary part_time female lag_m6 lag_u3 pp_lag_m6 pp_lag_u3 ofsted_ldr_2 ofsted_ldr_3 ofsted_ldr_4 [pweight=panel_wt2] if panel==1 & year==2015 , robust baseoutcome (0) rrr

[19] https://www.gov.uk/government/collections/ofsted-inspections-of-maintained-schools

[20] Based on total membership as reported on the TUC website of ATL (127.9k), NASUWT (294.2k), NUT (326.9k), UCAC (3.9k) and NAHT (28.5k) (total 777.5k), divided by total teaching FTE in the 2014 School Workforce Census (848.8k). ASCL and Voice membership were not reported by the TUC.

[21] School Workforce Census: https://www.gov.uk/government/publications/school-workforce-in-england-november-2012